



Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 244

Beta Regimes for the Yield Curve

Francesco Audrino and Enrico De Giorgi

May 2005

Beta Regimes for the Yield Curve*

Francesco Audrino[†]

University of Lugano

Enrico De Giorgi[‡]

University of Zurich and

University of Lugano

First Draft: 16th December 2004

This Version: 18th May 2005

Abstract

We propose an affine term structure model which accommodates non-linearities in the drift and volatility function of the short-term interest rate. Such non-linearities are a consequence of discrete beta-distributed regime shifts constructed on multiple thresholds. We derive iterative closed-form formula for the whole yield curve dynamics that can be estimated using a linearized Kalman filter. Fitting the model on US data, we collect empirical evidence of its potential in estimating conditional volatility and correlation across yields.

Keywords: Threshold Regime Switching Model, Affine Model, Term Structure of Interest Rate, Linearized Kalman Filter.

JEL Classification: E43, G12, C51, C52.

*Financial support from the Foundation for Research and Development of the University of Lugano is gratefully acknowledged. We received useful comments from Fabio Trojani.

[†]Institute of Finance, University of Lugano, via Buffi 13, CH-6900 Lugano, Switzerland, Tel. +41 +58 6664789, Fax. +41 +58 6664647, email: francesco.audrino@lu.unisi.ch.

[‡]Institute of Finance, University of Lugano, via Buffi 13, CH-6900 Lugano, Switzerland, Tel. +41 +58 6664677, Fax. +41 +58 6664647 and Institute for Empirical Research in Economics, University of Zurich, email: degiorgi@iew.unizh.ch, enrico.degiorgi@lu.unisi.ch.

1 Introduction

In the last years several studies collected empirical evidence that the univariate short-term interest rate dynamics over time can be accurately described using regime switching approaches; see, for example, Garcia and Perron (1996), Bansal and Zhou (2002) or Ang and Bekaert (2002a) in the Hamilton's Markovian regime switching framework, or Audrino (2004) in a threshold autoregressive regime switching framework. As Ang and Bekaert (2002b) showed, a parametric regime switching model can match well the non-linear short rate drift and volatility patterns present in real data and detected in several non-parametric studies of the term structure (see, for example, Ait-Sahalia 1996a, 1996b, or Stanton 1997). This finding gives a first reason about the importance of using models allowing for regime shifts when estimating the term structure and when pricing interest rate dependent instruments. In addition to that, shifts in conditional mean and volatility regimes impact in a economically significant way the whole term structure of interest rates and not only the short-term interest rate process. In fact, standard term structure models, like the Cox, Ingersoll, and Ross (1985) (CIR) model and like affine models that do not consider the possibility of discrete regime shifts, may lead to poor empirical performance when fitted to real data.

In this article, we propose a new term structure model where the short-term interest rate process as well as the market prices of risk are subjected to discrete regime shifts. A similar approach was proposed by Bansal and Zhou (2002) in the Hamilton's Markovian framework with constant transition probabilities across different linear processes. Generalizations of this model to incorporate time-varying transition probabilities were developed by Dai, Singleton, and Yang (2003) and Fink (2004). In such models, transition probabilities can be state-dependent (function of underlying factors or/and some exogenous macroeconomic variables), but the Markov process governing the regime changes is assumed to be independent from the factor process.

Following this direction and supported by empirical evidence, we also believe that the probability to be at any given time in a specific regime has to be time-varying and directly related to some relevant state variables. Nevertheless, our approach is different from the regime switching models introduced in the literature above. In our model the underlying regime variable does not follow a Markov process. Instead, regimes are directly characterized by multiple thresholds on the regime variable. As Audrino (2004) showed in his study, a threshold autoregressive regime switching model is able to estimate and forecast well the short-term interest rate dynamics. Moreover, in this model the relation between regimes and state variables, such as the short-term interest rate itself or other exogenous macroeconomic variables, is direct. As a consequence, regimes have a clear interpretation in terms of disjunct regions of the relevant state space and are determined by thresholds. Thus, the current regime can be derived and the model provides a forecasting tool. This is an improvement compared to other regime switching models presented in the literature. In fact, in the Hamilton's Markovian framework the current regime is not observed, but randomly distributed across the several possible regimes, according to a given probability distribution. Bansal and Zhou (2002), for example, identify the current regime assuming that it is the one that minimizes the pricing error on the estimates of the whole yield curve. In other words, the current regime is determined a-posteriori given the realization of the yield curve. This is clearly not a feasible strategy when the final goal is forecasting,

Starting from a one-period no arbitrage condition, we derive analytical iterative closed-form expressions for the whole term structure. We estimate our model on monthly US treasury bond yield data from the CRSP database using the linearized Kalman filter technique studies by Duffee and Stanton (2004), among others. The time period we consider goes from January 1960 to December 2001. Comparing the results

from our model with those from two alternative competitors we collect the following empirical evidence. First, our model suggests the presence of more than two regimes in the short-term interest rate process, consistently to previous results found in Audrino (2004). Second, when fitting our one-factor Beta model to US interest rate data we collect empirical evidence of its strong potential in estimating the yield curve, also in comparison to other single factor alternative approaches like the regime-switching specifications of Bansal and Zhou (2002) and the standard CIR model. In particular, our model shows minimal pricing errors (averaged on eleven different maturities) and is the only one that is able to mimic well different shapes of the realized yield curve. Thus, our one-factor Beta model joins at the the same time simplicity of the term structure dynamics with good empirical results

In the fourth section of the paper, we give some insight into how the single factor Beta model can be improved. We report some preliminary results for a two-factor generalization of the Beta model. In particular, the pricing errors of the two-factor Beta model are on average less then 20 basis points. This number is encouraging when compared, for example, to the 30 basis points showed on average by the two-factor CIR model.

The remainder of the paper is organized as follows. The Beta model for the term structure is introduced in Section 2. The results of our empirical analysis on US term structure data are summarized in Section 3. In Section 4 we give a short outlook of possible applications of our model as well as future developments. Section 5 concludes the paper.

2 The model

In this section we present our model for the state variable and we derive the no-arbitrage conditions for the term structure of interest rates. The model setup follows Bansal and Zhou (2002). We construct a Cox, Ingersoll, and Ross (1985) model (CIR model) for the state variable with regime shifts. The latter are assumed to be determined by a threshold model for the state variable and beta distributed dispersions around the current observation. We present below the theory for a one-factor model. The extension to more factors model gives rise to some theoretical open questions that are discussed in Section 4. In addition to that, note that the estimation of a two- or three-factors specification is clearly more computationally expensive since it involves an higher number of parameters.

Let $(X_t)_{t \geq 0}$ be a one-dimensional state variable and $(s_t)_{t \geq 0}$ be the discrete regime switching process with values $k \in \{1, \dots, K\}$. We assume that given the regime of the economy s_{t+1} at time $t+1$, the state variable satisfies the following dynamics (Cox, Ingersoll, and Ross 1985):

$$X_{t+1} - X_t = \kappa_{s_{t+1}} (\theta_{s_{t+1}} - X_t) + \sigma_{s_{t+1}} \sqrt{X_t} U_{t+1}, \quad (1)$$

where U_{t+1} are i.i.d. standard normally distributed innovations, and $\kappa_{s_{t+1}}$, $\theta_{s_{t+1}}$ and $\sigma_{s_{t+1}}$ are the regime-dependent mean reversion, long run mean and volatility parameters, respectively. We denote the regime-dependent conditional mean and conditional variance of the process given the information \mathcal{F}_t up to time t and the regime s_{t+1} by $\mu_{s_{t+1}}(X_t) = \kappa_{s_{t+1}} (\theta_{s_{t+1}} - X_t)$ and $\sigma_{s_{t+1}}^2(X_t) = \sigma_{s_{t+1}}^2 X_t$, respectively. Using the standard notation, we denote the difference $X_{t+1} - X_t$ by ΔX_{t+1} .

Contrary to the model proposed by Bansal and Zhou (2002), the regime switching process is determined

as

$$s_{t+1} = k \Leftrightarrow V_{t+1} \in (d_{k-1}, d_k] =: \mathcal{R}_k, \quad \text{for } k = 1, \dots, K, \quad (2)$$

where $d_0 < d_1 < \dots < d_K$ are thresholds for the state variable X_t and the real-valued random variable V_{t+1} satisfies

$$V_{t+1} | (X_t, s_t) \sim \text{Beta}(\alpha_{s_t}(X_t), \beta_{s_t}(X_t)), \quad (3)$$

i.e. V_{t+1} is conditionally beta distributed on $[d_0, d_K]$ with parameters $\alpha_{s_t}(X_t)$ and $\beta_{s_t}(X_t)$, given the observation of the state variable X_t and the regime s_t at time t .¹ Additionally, we initialize the process $V = (V_t)_{t \geq 0}$ imposing $V_0 = X_0$. Equations (2) and (3) explain the following dynamics for the regime process: given the observation of the state variable X_t at time t , the time $t + 1$ regime is determined by a conditionally beta distributed variable, where the parameters of the distribution are functions of X_t . In other words, the probability that next regime occurs depends on both the current observation of the state variable and a fixed threshold structure for the state variable. This implies one of the main property of the model: given the threshold structure and the current observation of the state variable, we are able to derive the current regime. This is not the case in the Markovian framework of Bansal and Zhou (2002). In their model, the current regime is not observed but randomly distributed between the possible regimes with a given probability distribution. For this reason, in Bansal and Zhou (2002) the current regime is determined a-posteriori such that it minimizes the pricing error on the estimates of the whole term structure. As a consequence, their model definitively depends not only on the realized state variable, but also on the whole realized term structure. On the contrary, our model only depends on the realized state variable.

The functions $\alpha_{s_t} : \mathbb{R} \rightarrow \mathbb{R}^+$ and $\beta_{s_t} : \mathbb{R} \rightarrow \mathbb{R}^+$ are assumed to be continuously differentiable. If both α_{s_t} and β_{s_t} are constant, then we have constant transition probabilities. Nevertheless, the current regime can still be derived from the threshold structure. Later, we will impose some restrictions on α_{s_t} and β_{s_t} . In particular, we will assume that for all X_t and all s_t , the conditional mode of V_{t+1} corresponds to X_t , i.e.

$$X_t = d_0 + (d_K - d_0) \frac{\alpha_{s_t}(X_t) - 1}{\alpha_{s_t}(X_t) + \beta_{s_t}(X_t) - 2} \Leftrightarrow \alpha_{s_t}(X_t) = \frac{(2 - \beta_{s_t}(X_t)) \tilde{X}_t - 1}{\tilde{X}_t - 1}, \quad (4)$$

where $\tilde{X}_t = \frac{X_t - d_0}{d_K - d_0}$ is the shifted and normalized state variable. In this case, equation (2) implies that given X_t and s_t , depending on α_{s_t} and β_{s_t} , the process V_{t+1} determining the regime at time $t + 1$ is distributed around the current observation of the state variable. In fact, the restriction (4) for the parameters α_{s_t} and β_{s_t} is supported by empirical evidence collected in the literature. For example, Bansal and Zhou (2002) found that the probability of staying in the same regime is very high (about 98%). Assuming (4) we ensure the beta transition probabilities (3) in our model to be concentrated around the current value of the state variable.

Above we denoted by $\mathcal{F}_t = \sigma(X_u, s_u : u \leq t)$ the information available up to time t . In our model it corresponds to the σ -algebra generated by $(X_u)_{u \leq t}$ and $(s_u)_{u \leq t}$ meaning that the state variable and the regime process up to time t are observable at time t . The conditional cumulative function of X_{t+1} given

\mathcal{F}_t can be computed as

$$\begin{aligned}
F_t(x) &= \mathbb{P}[X_{t+1} \leq x | \mathcal{F}_t] \\
&= \sum_{k=1}^K \mathbb{P}[X_{t+1} \leq x | X_t, s_t, V_{t+1} \in (d_{k-1}, d_k]] \mathbb{P}[V_{t+1} \in (d_{k-1}, d_k] | X_t, s_t] \\
&= \sum_{k=1}^K \Phi\left(\frac{x - X_t - \kappa_k(\theta_k - X_t)}{\sigma_k \sqrt{X_t}}\right) \left[I(\tilde{d}_k; \alpha_{s_t}(X_t), \beta_{s_t}(X_t)) - I(\tilde{d}_{k-1}; \alpha_{s_t}(X_t), \beta_{s_t}(X_t)) \right]
\end{aligned}$$

where $\tilde{d}_k = \frac{d_k - d_0}{d_K - d_0}$, $k = 1, \dots, K$, are the normalized thresholds, and $I(x; \alpha, \beta)$ is the incomplete beta function

$$I(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^x (1-u)^{\beta-1} u^{\alpha-1} du.$$

Φ denotes as usual the standard normal distribution function. The conditional density function of X_{t+1} given \mathcal{F}_t follows directly, equals

$$f_t(x) = \sum_{k=1}^K \frac{1}{\sigma_k \sqrt{X_t}} \varphi\left(\frac{x - X_t - \kappa_k(\theta_k - X_t)}{\sigma_k \sqrt{X_t}}\right) \left[I(\tilde{d}_k; \alpha_{s_t}(X_t), \beta_{s_t}(X_t)) - I(\tilde{d}_{k-1}; \alpha_{s_t}(X_t), \beta_{s_t}(X_t)) \right] \quad (5)$$

and corresponds to a weighted sum of normal densities. φ in (5) denotes the standard normal density function. In the above expression for the conditional density, the last factor of each term is the conditional transition probability. In fact, let $p_k(X_t) = \mathbb{P}[s_{t+1} = k | \mathcal{F}_t]$ be the conditional transition probability to regime k given \mathcal{F}_t , then from equations (1) and (2) we have:

$$p_k(X_t) = \mathbb{P}[V_{t+1} \in (d_{k-1}, d_k] | \mathcal{F}_t] = I(\tilde{d}_k; \alpha_{s_t}(X_t), \beta_{s_t}(X_t)) - I(\tilde{d}_{k-1}; \alpha_{s_t}(X_t), \beta_{s_t}(X_t)). \quad (6)$$

For β_{s_t} constant or linear in X_t and assuming that the mode of the Beta distribution corresponds to X_t , the function p_k are obviously non-linear in X_t . This implies that our model is able to produce non-linear conditional means and volatilities. In fact, nowadays there is strong empirical evidence that the realized short-term interest rate process shows non-linearities in the first and second conditional moments (see for instance Aït-Sahalia (1996a), Stanton (1997) and Ang and Bekaert (2002b)). In particular, in our model the conditional mean $\mu_{t+1}(X_t) = \mathbb{E}[\Delta X_{t+1} | \mathcal{F}_t]$ and the conditional variance $\sigma_{t+1}^2(X_t) = \text{Var}[\Delta X_{t+1} | \mathcal{F}_t]$ of the process ΔX_{t+1} given \mathcal{F}_t can be computed as

$$\begin{aligned}
\mu_{t+1}(X_t) &= \sum_{k=1}^K \mu_k(X_t) p_k(X_t) = \sum_{k=1}^K \kappa_k(\theta_k - X_t) p_k(X_t), \quad \text{and} \\
\sigma_{t+1}^2(X_t) &= \sum_{k=1}^K \sigma_k^2 X_t p_k(X_t) + \sum_{k=1}^K (\mu_{t+1} - \mu_k(X_t))^2 p_k(X_t).
\end{aligned}$$

These are clearly non-linear functions of the current realization X_t . The shapes of these functions are shown in Figures 1 and 2.

[Figure 1 about here.]

[Figure 2 about here.]

They show similar patterns to those reported in the literature using other parametric and non-parametric approaches (see, for example, Ang and Bekaert 2002b).

To finish this section, we derive the no arbitrage conditions imposed by our model on the affine price process of the zero-coupon bond. Let $P(t, n)$ be the price of a zero-coupon bond issued at time t with maturity n . We assume that $P(t, n)$ is an affine function of X_t , i.e. we find regime dependent parameters $A_{s_t}(n)$ and $B_{s_t}(n)$ such that

$$P(t, n) = \exp \{ -A_{s_t}(n) - B_{s_t}(n) X_t \}. \quad (7)$$

As in Bansal and Zhou (2002), we impose that $A_{s_t}(0) = B_{s_t}(0) = 0$ and $A_{s_t}(1) = 0$, $B_{s_t}(1) = 1$. The latter assumption implies that $r_{f,t} = X_t$, i.e. the state variable X_t corresponds to the short interest rate. Let

$$h(t+1, n) = \frac{P(t+1, n-1)}{P(t, n)}$$

be the one-period gross return at time $t+1$. Moreover, let

$$M_{t,t+1} = \exp \left[-r_{f,t} - \left(\frac{\lambda_{s_{t+1}}}{\sigma_{s_{t+1}}} \right)^2 \frac{X_t}{2} - \frac{\lambda_{s_{t+1}}}{\sigma_{s_{t+1}}} \sqrt{X_t} U_{t+1} \right] \quad (8)$$

be the one-period pricing kernel at time t . Then the one-period no-arbitrage condition for $(t, t+1]$ is

$$\mathbb{E} [M_{t,t+1} P(t+1, n-1) | \mathcal{F}_t] = P(t, n)$$

or, equivalently,

$$\mathbb{E} [M_{t,t+1} h(t+1, n) | \mathcal{F}_t] = 1.$$

The one-period no-arbitrage condition states that the bond price $P(t, n)$ at time t corresponds to the discounted bond price $P(t+1, n-1)$ at time $t+1$, conditioned on the current information. Therefore, conditioning on the regimes, we obtain

$$\begin{aligned} 1 &= \sum_{k=1}^K \mathbb{E} [\mathbb{E} [M_{t,t+1} h(t+1, n) | \mathcal{F}_t, s_{t+1} = k] | \mathcal{F}_t] \mathbb{P} [s_{t+1} = k | \mathcal{F}_t] \\ &= \sum_{k=1}^K \mathbb{E} [\mathbb{E} [M_{t,t+1} h(t+1, n) | \mathcal{F}_t, s_{t+1} = k] | \mathcal{F}_t] p_k(X_t) \end{aligned}$$

Given \mathcal{F}_t and $s_{t+1} = k$, the random variable $M_{t,t+1} h_{n,t+1}$ is log-normal distributed with mean

$$\exp \left(\mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \right),$$

where

$$\mu_k(t, n) = -r_{f,t} - \left(\frac{\lambda_k}{\sigma_k} \right)^2 \frac{X_t}{2} - A_k(n-1) - B_k(n-1) \mathbb{E} [X_{t+1} | \mathcal{F}_t, s_{t+1} = k] + A_{s_t}(n) + B_{s_t}(n) X_t,$$

$$\sigma_k(t, n)^2 = \left(B_k(n-1) \sigma_k + \frac{\lambda_k}{\sigma_k} \right)^2 X_t = B_k(n-1)^2 \sigma_k^2 X_t + 2 \lambda_k B_k(n-1) X_t + \left(\frac{\lambda_k}{\sigma_k} \right)^2 X_t.$$

Consequently, from the no-arbitrage condition we derive that

$$\begin{aligned} 1 &= \sum_{k=1}^K \exp \left(\mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \right) p_k(X_t) \\ &= \frac{1}{B(\alpha_{s_t}(X_t), \beta_{s_t}(X_t))} \sum_{k=1}^K \int_{\tilde{d}_{k-1}}^{\tilde{d}_k} \exp \left(\mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \right) (1-u)^{\beta_{s_t}(X_t)-1} u^{\alpha_{s_t}(X_t)-1} du. \end{aligned}$$

The last equation can be equivalently rewritten as

$$0 = \sum_{k=1}^K \left[\exp \left(\mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \right) - 1 \right] \int_{\tilde{d}_{k-1}}^{\tilde{d}_k} (1-u)^{\beta_{s_t}(X_t)-1} u^{\alpha_{s_t}(X_t)-1} du.$$

Similarly to Bansal and Zhou (2002), we solve this equation using an approximation. In particular, we derive the first order approximation of the latter expression for $X_t \approx 0$. Let

$$g(x; u) = (1-u)^{\beta(x)-1} u^{\alpha(x)-1}.$$

Then

$$\partial_x g(x; u) = (1-u)^{\beta(x)-1} u^{\alpha(x)-1} [\beta'(x) \log(1-u) + \alpha'(x) \log(u)]$$

and the first-order approximation in $x \approx 0$ is

$$g(x; u) \approx (1-u)^{\beta(0)-1} u^{\alpha(0)-1} [1 + (\beta'(0) \log(1-u) + \alpha'(0) \log(u)) x].$$

Moreover, for $x \approx 0$, $\exp(x) - 1 \approx x$. Note that this approximation has sense since in our model x is the value of the short rate (in our sample, the short rate is always less than 17%). Thus, the first order no-arbitrage condition is given by

$$0 \approx \sum_{k=1}^K \left(\mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \right) \int_{\tilde{d}_{k-1}}^{\tilde{d}_k} (1-u)^{\beta_{s_t}(0)-1} u^{\alpha_{s_t}(0)-1} [1 + (\beta'_{s_t}(0) \log(1-u) + \alpha'_{s_t}(0) \log(u)) X_t] du.$$

When dividing the last equation by $B(\alpha_{s_t}(0), \beta_{s_t}(0))$ and defining the following quantities

$$\begin{aligned} p_{s_t k} &= \frac{1}{B(\alpha_{s_t}(0), \beta_{s_t}(0))} \int_{\tilde{d}_{k-1}}^{\tilde{d}_k} (1-u)^{\beta_{s_t}(0)-1} u^{\alpha_{s_t}(0)-1} du \\ &= I(\tilde{d}_k; \alpha_{s_t}(0), \beta_{s_t}(0)) - I(\tilde{d}_{k-1}; \alpha_{s_t}(0), \beta_{s_t}(0)) \end{aligned}$$

and

$$q_{s_t k} = \frac{1}{B(\alpha_{s_t}(0), \beta_{s_t}(0))} \int_{\tilde{d}_{k-1}}^{\tilde{d}_k} (1-u)^{\beta_{s_t}(0)-1} u^{\alpha_{s_t}(0)-1} (\beta'_{s_t}(0) \log(1-u) + \alpha'_{s_t}(0) \log(u)) du,$$

we get that

$$0 \approx \sum_{k=1}^K \left(\mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \right) (p_{s_t k} + q_{s_t k} X_t).$$

Note that from the definition of $p_{s_t k}$ follows that for all s_t , $\sum_{k=1}^K p_{s_t k} = 1$. Moreover, $q_{s_t k} = 0$ if α_{s_t} and β_{s_t} are both constant, i.e. do not depend on the current realization of the state variable.

Finally, since

$$\begin{aligned} & \mu_k(t, n) + \frac{1}{2} \sigma_k(t, n)^2 \\ &= -X_t - A_k(n-1) - B_k(n-1) (X_t + \kappa_k (\theta_k - X_t)) \\ & \quad + A_{s_t}(n) + B_{s_t}(n) X_t + \frac{1}{2} B_k(n-1)^2 \sigma_k^2 X_t + \lambda_k B_k(n-1) X_t \\ &= - \left[1 + (1 - \kappa_k - \lambda_k) B_k(n-1) - \frac{1}{2} B_k(n-1)^2 \sigma_k^2 - B_{s_t}(n) \right] X_t \\ & \quad - [A_k(n-1) + \kappa_k \theta_k B_k(n-1) - A_{s_t}(n)] \end{aligned}$$

the first order no-arbitrage conditions become

$$\begin{aligned} A_{s_t}(n) &= \sum_{k=1}^K [A_k(n-1) + \kappa_k \theta_k B_k(n-1)] p_{s_t k} \\ B_{s_t}(n) &= \sum_{k=1}^K [A_k(n-1) + \kappa_k \theta_k B_k(n-1)] \left(q_{s_t k} - p_{s_t k} \sum_{l=1}^K q_{s_t l} \right) \\ & \quad + \sum_{k=1}^K \left[1 + (1 - \kappa_k - \lambda_k) B_k(n-1) - \frac{1}{2} B_k(n-1)^2 \sigma_k^2 \right] p_{s_t k}. \end{aligned}$$

Compared to Bansal and Zhou (2002), the iterations for the parameters of the affine model (7) under our model structure for the state variable present the additional term

$$\sum_{k=1}^K [A_k(n-1) + \kappa_k \theta_k B_k(n-1)] \left(q_{s_t k} - p_{s_t k} \sum_{l=1}^K q_{s_t l} \right)$$

that enters in the calculation of B_{s_t} . This term relates to the first order approximation of the transition probabilities and obviously vanishes under the assumption of constant transition probabilities made by Bansal and Zhou (2002).

When rewriting in a vector notation the iterations implied by the above derived no-arbitrage conditions, we get the following. Let $A(n) = (A_1(n), \dots, A_K(n))'$ and $B(n) = (B_1(n), \dots, B_K(n))'$. Moreover, let we define $P = (p_{lk})_{l,k=1,\dots,K}$, $Q = (q_{lk})_{l,k=1,\dots,K}$ and, for all $k = 1, \dots, K$

$$\begin{aligned} B_k^1(n) &= \kappa_k \theta_k B_k(n), \\ B_k^2(n) &= 1 + (1 - \kappa_k - \lambda_k) B_k(n-1) - \frac{1}{2} B_k(n-1)^2 \sigma_k^2, \end{aligned}$$

$B^i(n) = (B_1^i(n), \dots, B_K^i(n))'$ for $i = 1, 2$. Then

$$A(n) = P [A(n-1) + B^1(n-1)], \quad (9)$$

$$B(n) = [Q - P \otimes Q\mathbf{1}] [A(n-1) + B^1(n-1)] + PB^2(n-1) \quad (10)$$

$$A(0) = B(0) = 0. \quad (11)$$

Finally, let $Y(t, n)$ be the yield of the zero-coupon bond with maturity n , i.e.

$$P(t, n) = e^{-nY(t, n)}.$$

Then, from equation (7) we obtain that

$$Y(t, n) = \frac{A_{s_t}(n)}{n} + \frac{B_{s_t}(n)}{n} X_t. \quad (12)$$

3 Empirical analysis

3.1 Data description

The data set used in this work is obtained from the Center for Research in Security Prices (CRSP) and have being first constructed by Fama (1984) and Fama and Bliss (1987). The data consists of 504 monthly observations of annualized discount bond yields for the period from January 1960 to December 2001, with maturities 1, 2, 3, 4, 5, 6 months and 1, 2, 3, 4, 5 years.² The same (or part of the same) data set was used by Bansal and Zhou (2002) and Longstaff and Schwarz (1992), among others, and is quite standard in the term structure literature.

In Table 1 we summarize the main statistics of the data.

[Table 1 about here.]

As expected, the empirical distribution of discount bond yields are leptokurtic and positively skewed. The standard deviation, skewness and kurtosis are systematically higher for short maturities than for the longer ones, while the sample mean shows the opposite behavior (i.e. the average yield curve is upward sloping).

Figure 3 shows the term structure dynamics over the sample period from January 1960 to December 2001.

[Figure 3 about here.]

We observe several dramatic changes of the interest rates. In particular, interest rates increase sharply and show considerable higher volatility than average during the OPEC oil crises 1973-1975, the Federal Reserve monetary experiment from October 1979 to October 1982 of targeting non-borrowed reserves instead of interest rates (Friedman 1984) and after the market crash of October 1987.

Figure 4 gives qualitatively distinct shapes of the yield curve present in our sample period: concave, downward humped, upward humped, convex. The average yield curve is upward sloping and concave.

[Figure 4 about here.]

3.2 Preliminary short rate Beta model estimation

We estimate the Beta model following a two-steps procedure. First, we use pseudo maximum-likelihood estimation to fit the model to the data of the short-rate, using the 1-month yield data. The conditional density function is given in equation (5). This first step identifies the number of regimes and the threshold structure. Second, we apply the extended Kalman filter estimation on the data of the whole term structure, where the threshold structure and the starting parameters of the Kalman filter recursion are the optimal ones from the first step. Regime-dependent market prices of risk are initially set equal to zero. In this section we focus on the estimation of the regimes (thresholds) for the short-rate process.

We estimate two different specifications of the Beta model. In both specifications we assume that the parameters $\alpha_{s_t}(X_t)$ and $\beta_{s_t}(X_t)$ of the Beta distribution determining the next regime satisfy equation (4). This assumption is equivalent to imposing the conditional mode of the Beta distribution being equal to the current observation of the state variable, as discussed in the previous section. Given equation (4) we choose two functional forms for the β_{s_t} functions. In particular, we impose a parametric model for the β_{s_t} , where we allow the β_{s_t} to be constant or linear functions of the state variable. In fact, since we take first order approximations of the single period no-arbitrage condition, more complexity is not needed (see the previous section).

The optimal functional form of the β_{s_t} and the optimal threshold structure (and, consequently, the number of regimes) are estimated using a binary tree construction. This procedure follows closely the one introduced by Audrino and Bühlmann (2001). Summarizing, the estimation procedure is as follows.

- (a) The minimizing criterion to select the best threshold is always the negative log-likelihood based on the conditional density (5). The parametric form of β_{s_t} in (4) is always constant or linear.
- (b) Optimization with threshold functions in item (a) becomes an estimation and model selection problem. The former is done by pseudo maximum likelihood. For the latter, an exhaustive search is computationally prohibitive. Like in Audrino and Bühlmann (2001) we propose a tree-structured partial search: within a data-determined tree structure, the optimal model is estimated using the AIC criterion. In particular, we choose as possible threshold candidates 1/8-quantiles of the sample short-rate observations.

One of the main advantages of this approach is that the threshold structure and the optimal number of regimes is determined endogenously in the estimation. This is not the case, for example, in a standard Markovian regime-switching model where the optimal number of regimes is given a-priori (and is in general low, due to the high computational complexity involved in the estimation). For all details about the estimation procedure we refer to Audrino (2004).

The maximum-likelihood estimates of the short-rate model using the above mentioned strategy are shown in Table 2.

[Table 2 about here.]

We mainly focus our discussion here on the estimated threshold structure and β_{s_t} functions. The other parameters display the same patterns as those reported in the literature for similar models for the short-term interest rate process (see Table 2, Panel B). Moreover, a detailed discussion of this preliminary parameter estimates for the conditional mean and variance dynamics are far away from the scope of this

paper. In the next section, we are going to discuss in a more exhaustive way the Kalman filter estimates for regime-dependent conditional means and volatilities based on the whole term structure data.

We estimated several Beta model specifications. Table 2 only reports the parameter estimates for the best fits, i.e. the two regimes Beta model with linear β_{s_t} and the three regimes Beta model with constant β_{s_t} . Goodness of fit results for the estimation of our Beta model specification, as well as other standard approaches used in the literature are reported in Table 3.

[Table 3 about here.]

In particular, the model with three regimes and constant beta parameters (later called 3-regimes constant Beta model) is selected by *AIC*. This model also minimizes the estimation error (measured by means of the mean absolute error (MAE) statistic and by means of the root mean squared error (RMSE) statistic) for both the drift and the volatility. Consequently, also within this model setup, similar to Audrino (2004), we obtain empirical support to the existence of more than two regimes for the short-rate process. Moreover, when computing the standard errors of the parameters using a classical bootstrap based methodology (see, for example, Efron and Tibshirani 1993) we find also statistical support for the presence of more than two drift and volatility regimes. The optimal thresholds values are 3.859% and 8.382%. The latter also corresponds to the optimal threshold in the 2-regimes linear Beta model. Directly from the construction of our binary tree, we identify a strict relationship between different regimes for the short-rate and the level of the short-rate itself. In fact, the estimated first threshold corresponds to the 25% quantile, while the estimated second threshold is the 87.5% quantile. That means that dates with corresponding one-period behind short-term interest rates that lie below the 3.859% threshold are characterized by conditional means and variances following the first regime dynamics. Similarly, dates with corresponding one-period behind short-term interest rates that are between the 3.859% and the 8.382% thresholds have conditional first and second moments following the second regime dynamics. Finally, dates with one-period behind short-term interest rates above the 8.382% threshold are associated with the third regime. For our data sample, we lie most of the time in the middle regime.

The level of the short rate also impacts the transition probabilities. The coefficients of the β_{s_t} functions are reported in Table 2 (Panel A). The resulting conditional density functions are also plotted in Figure 5 for the 3-regimes constant Beta model. Note that the fact that the β_{s_t} functions are constant does not mean that the Beta transition probabilities do not depend on the level of the short rate. In fact, imposing (4) we have that the α_{s_t} functions always depend on X_t except in the case where $\beta_{s_t} \equiv 1$.

[Figure 5 about here.]

The upper panel corresponds to the first regime: the probability of shifting to higher regimes is high. The conditional density function is positively skewed with right fat tail. The middle panel shows the conditional density functions for the second regime. These are symmetric around the current observation and less dispersed compared to the first regime. This means that lying in the middle regime, we have a high probability to stay in the same regime. Finally, the lower panel shows the density function for the third regime. The shape is similar to that of the second regime, but the conditional density function are much more concentrated around the current observations, in particular for very high values of the short-rate. The shape of the conditional density functions for the second and third regimes suggests that higher regimes are more likely to persist.

To end this section, we compare the goodness of fit results from the preliminary estimated Beta model with those from (i) a global CIR model, (ii) the regime-switching model of Bansal and Zhou

(2002) with two-regimes and constant transition probabilities, and (iii) with a generalization of the two-regimes model of Bansal and Zhou (2002), where transition probabilities are allowed to depend from the current level of the short rate (see Gray 1996). The goodness of fit statistics are reported in Table 3. In addition, the estimated conditional variances and the conditional probabilities of transition for the regime-switching model with time-varying transition probabilities and the Beta models are plotted in Figures 6-8, respectively.

[Figure 6 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

As expected, all the models considered in the Figures are able to mimic well periods of high and low volatilities and short-term interest rates. Nevertheless, the 3-regimes constant Beta model clearly outperforms the other approaches with respect to several estimation measures reported in Table 3. In particular, within the class of models considered, the 3-regimes constant Beta model shows minimal mean absolute error (MAE) and minimal root mean squared error (RMSE) for the drift. The latter and the R^2 (see Gray 1996) are also minimal for the volatility.

3.3 Term structure model estimation

The second step of the estimation procedure is based on a Kalman filter estimation for the data of the whole term structure. In this second step we keep fixed the structure of the regimes (i.e. the optimal thresholds for the state variable) estimated in the first step. We restrict ourself to the 3-regimes constant Beta model from the previous subsection since it yields the best estimation results for the conditional first and second moment dynamics of the short-rate process.

The theoretical and empirical properties of a linearized Kalman filter estimation applied to different term structure models was studied among others by Duffee and Stanton (2004). In their work, Duffee and Stanton argue in favor of the Kalman filter technique when compared to the Efficient Method of Moments of Gallant and Tauchen (1996). This despite the asymptotic equivalence of the latter to a maximum-likelihood estimation.

The next paragraph briefly describes the extended Kalman filter estimation technique. We refer to Duffee and Stanton (2004) for a detailed discussion.

Let us consider the average state variable dynamics (in our one-factor Beta model specification identified with the short-term interest rate), i.e.

$$X_{t+1} = X_t + \mu_{t+1}(X_t) + \sigma_{t+1}(X_t) U_{t+1},$$

where

$$\begin{aligned} \mu_{t+1}(X_t) &= \sum_{k=1}^K \kappa_k \theta_k p_k(X_t) - X_t \sum_{k=1}^K \kappa_k p_k(X_t), \quad \text{and} \\ \sigma_{t+1}^2(X_t) &= \sum_{k=1}^K \sigma_k^2 p_k(X_t) + \sum_{k=1}^K (\mu_{t+1}(X_t) - \mu_k(X_t))^2 p_k(X_t). \end{aligned}$$

We rewrite these dynamics as follows:

$$X_{t+1} = F_0(X_t, \rho) + F_1(X_t, \rho) X_t + \tilde{U}_{t+1} \quad (13)$$

where $\rho = \{\beta_k, \kappa_k, \theta_k, \sigma_k, \lambda_k, k = 1, \dots, K\}$ is the vector of the parameters that have to be estimated (K being the number of regimes), $F_0(X_t, \rho) = \sum_{k=1}^K \kappa_k \theta_k p_k(X_t)$, $F_1(X_t, \rho) = 1 - \sum_{k=1}^K \kappa_k p_k(X_t)$ and \tilde{U}_{t+1} is a Gaussian distributed innovation variable with mean 0 and variance $Q(X_t, \rho) = \sigma_{t+1}^2(X_t)$. Let we denote by $\mathbf{Y}_t = (Y(t, n_1), \dots, Y(t, n_\tau))'$ the yields at the different maturities present in our data sample. Then, by equation (12) we have that

$$\mathbf{Y}_t = H_0(X_t, \rho) + H_1(X_t, \rho) X_t + \epsilon_t \quad (14)$$

where H_0 and H_1 are n_τ -dimensional functions of X_t and ρ , that are defined by

$$H_{0,j}(X_t, \rho) = \frac{1}{n_j} \sum_{k=1}^K A_k(n_j) 1_{X_t \in (d_{k-1}, d_k]}, \quad H_{1,j}(X_t, \rho) = \frac{1}{n_j} \sum_{k=1}^K B_k(n_j) 1_{X_t \in (d_{k-1}, d_k]}$$

for $j = 1, \dots, \tau$ and, ϵ_t is an additional multivariate distributed idiosyncratic noise term with mean zero and variance-covariance matrix $R(\epsilon) = \epsilon Id$.³ As already said at the beginning of this section, we keep fixed the threshold structure determined in the first step of the estimation procedure as described in section 3.2.

Equations (13) and (14) are the linear models for the term structure. The functions F_0 , F_1 , H_0 and H_1 depend on both the current observation X_t and the parameters ρ . The Kalman filter recursion starts with a vector of parameters ρ_0 . In our estimation procedure a reasonable starting value for ρ is given by the set of maximum-likelihood parameter estimates from the previous subsection (estimated on the short-term interest rate process, see Table 2, Panel B). In addition to that, we initialize the regime-dependent market prices of risk to zero. Using the parameter ρ_0 we simulate the unconditional average and unconditional variance of X_{t+1} and we use these values as initial proxy for the mean and variance, respectively. We denote these quantities by $x_{0|0}^P$ and $P_{0|0}$. The extended Kalman filter recursion works as follows. Let $(\mathbf{y}_t)_{t=1, \dots, T}$ be the vector of observed yields at time t , then for $\rho = \rho_0$ and $t = 0, \dots, T-1$:

- (i) Let $F_{0t} = F_0(x_{t|t}, \rho)$, $F_{1t} = F_1(x_{t|t}, \rho)$ and $Q_t = Q(x_{t|t}, \rho)$. Compute the one-period-ahead predictions for the state variable $x_{t+1|t}^P = F_{0t} + F_{1t} x_{t|t}^P$, $P_{t+1|t} = F_{1t}' P_{t|t} F_{1t} + Q_t$,
- (ii) Let $H_{0t} = H_0(x_{t+1|t}^P, \rho)$, $H_{1t} = H_1(x_{t+1|t}^P, \rho)$. Compute the one-period ahead predictions for the yields $\mathbf{y}_{t+1|t}^P = H_{0t} + H_{1t} x_{t+1|t}^P$, $V_{t+1|t} = H_{1t} P_{t+1|t} H_{1t}' + R$.
- (iii) Compute the estimation error $\mathbf{e}_{t+1} = \mathbf{y}_{t+1} - \mathbf{y}_{t+1|t}^P$.
- (iv) Approximate the $(t+1)$ -period log-likelihood by

$$f_{t+1} = -\frac{1}{2} \left[\dim(\mathbf{y}_{t+1}) \log(2\pi) + \log |V_{t+1|t}| + \mathbf{e}_{t+1}' V_{t+1|t}^{-1} \mathbf{e}_{t+1} \right]$$

- (v) Update the prediction of the state variable: $x_{t+1|t+1}^P = x_{t+1|t}^P + P_{t+1|t} H_{1t}' V_{t+1|t}^{-1} \mathbf{e}_{t+1}$ and $P_{t+1|t+1} = P_{t+1|t} - P_{t+1|t} H_{1t}' V_{t+1|t}^{-1} H_{1t} P_{t+1|t}$.

The Kalman filter estimate solves

$$\hat{\rho} = \arg \max_{\rho} \sum_{t=1}^T f_t.$$

The Kalman filter estimate for the single factor 3-regimes constant Beta model are reported in Table 4.

[Table 4 about here.]

All the parameters, except for the speed of reversion parameter in the first regime are significant at the 1% confidence level rejecting the null hypothesis of the parameters to be equal to zero. The level of mean reversion parameter of the second and third regime belongs to the corresponding interval of the threshold model for the short rate, while the same parameter lies outside of the interval for the first regime. This latter regime presents a negative long-run mean and a negative speed of reversion. This implies that conditioned on the next regime being the first regimes, the drift of the short rate is strictly positive, i.e. on average the short rate increases when the next regime is the first regime. This dynamics is also supported by the estimated value $\beta_{1,0}$ of the Beta distribution for the first regime that is reported in Table 4. In fact, the estimated $\beta_{1,0}$ parameter implies that the probability of shifting away from the first regime to higher regimes is high (see, once again, Figure 5, upper panel). Note that the estimated regime-dependent constant beta functions are fairly equal to those initially estimated on the short rate process (see also the discussion after 5). On the other side, when the next regime are the second or the third one, the short rate tends to revert to the respective regime-dependent long-run mean.

The first regime (that correspond to small level for the short rate) has volatility parameters that are larger than those of the second, while the last regime presents the smallest volatility parameters. The first regime represents therefore the more volatile regime, while the third regime is far away the less volatile. At the same time, we estimate a negative market price of risk in the first regime, whereas it is positive in the other two regimes. Thus, contrary to Bansal and Zhou (2002) we find that the most volatile regime display a negative market price of risk. This result may be due to the fact that the more risk we take during periods characterized by regime one does not pay since market is in general decreasing. On the contrary, the second and third regimes are similar to those found by Bansal and Zhou (2002), the former being characterized by both high volatility and market price of risk. Thus, the difference between our results and those in Bansal and Zhou (2002) may be a direct consequence of the fact that in our model (i) we have a regime more and (ii) regimes are characterized using a threshold structure for the short rate.

The estimated yield curves from the constant Beta model superimposed on the realized ones for some given dates in our sample period are displayed in Figure 9.

[Figure 9 about here.]

Dates in the sample period are chosen such that yield curve estimates are computed from all the three regimes present in the model (i.e. for almost all the range of realized values of the short rate). Figure 9 shows well the ability of the one-factor 3-regimes constant Beta model to mimic different shapes of the realized yield curves. As it has been reported in the literature by several empirical studies, this is generally not true for other one-factor affine models of the term structure that only provide satisfactory results when extended to their multiple-factors versions (see Figure 10).

To end the empirical analysis, we compare the pricing errors implied by our model estimation to those from other model fits. In particular, we compare the results from the one-factor 3-regimes constant Beta model to those from the (i) one factor CIR model and (ii) one factor two-regimes specification of Bansal

and Zhou (2002). Results for the average absolute pricing error are summarized in Table 5. Pricing errors are averaged over the whole yield curve. In particular, in our case pricing errors are computed over ten different maturities. In the last columns of the table we also report results for two-factors specification of the above mentioned models. The discussion of these results is left to the following section.

[Table 5 about here.]

The 3-regimes constant Beta model outperforms the two other model specifications in term of the mean average absolute pricing errors. Moreover, the variance of the absolute pricing errors as well as the maximal pricing error are clearly below those from both the classical CIR and the regime-switching model of Bansal and Zhou (2002). This result confirms and supports the graphical conclusion that we reached after Figure 9.

4 Generalizations and possible applications

The regime-switching model presented in this work represents a simple but at the same time flexible tool for modelling and forecasting the yield curve. The results obtained for the one-factor 3-regimes Beta model are very encouraging and provide a solid basis for future research, concerning both theoretical and empirical issues.

In this section we briefly discuss possible generalizations of the Beta model. First, we present an extension of the one-factor model to multiple factors model. We also provide preliminary results of a two-factors Beta model. Second, we mention how to incorporate exogenous macroeconomic variables in the regime switching process. This can be crucial for improving estimates and forecasts of the yield curve dynamics. Finally, we mention possible applications of the model to the pricing of fixed income instruments and to interest rate risk management.

Two or multiple factors model. The one-factor affine model (7) can be easily extended in order to include multiple factors. Let X_1, \dots, X_M be a collection of state variables, where each factor X_m follows the regime switching CIR specification (1). Moreover, let suppose that the risk-free rate of return is the sum of the M factors $r_{f,t} = X_{1t} + \dots + X_{Mt}$ and the bond price satisfies the following multiple factors affine model:

$$P(t, n) = \exp \left\{ - \sum_{m=1}^M A_{m,s_{mt}}(n) - \sum_{m=1}^M B_{m,s_{mt}}(n) X_{mt} \right\}.$$

Therefore, the yield of a zero-coupon bond with time-to-maturity n is given by

$$Y(t, n) = \sum_{m=1}^M \frac{A_{m,s_{mt}}(n)}{n} + \sum_{m=1}^M \frac{B_{m,s_{mt}}(n)}{n} X_{mt}.$$

The coefficients $A_{m,s_{mt}}(n)$ and $B_{m,s_{mt}}(n)$ depend on the current regime s_{mt} for the state variable X_m . The pricing kernel can be easily obtained from the one-factor pricing kernel of equation (8), where the univariate risk premium is replaced by the sum of M square-root processes X_1, \dots, X_M . Similarly to the one-factor case, the next period regime for X_m is determined by a conditional Beta distributed random variable V_{mt} given X_{mt} and a threshold structure for X_m . Consequently and differently from the simplifying

assumption made in Bansal and Zhou (2002), we cannot take the same regime switching process for all factors. Moreover, conditioned on the current observation of the state variables, the regime switching processes might be correlated. In this case, the derivation of the affine model for the term structure is more complicated, since the probability of transitions depend on the multivariate conditional distribution. However, under the assumption of independent state variables, the first order one-period no-arbitrage condition can be simplified and the coefficients $A_{m,s_{mt}}(n)$ and $B_{m,s_{mt}}(n)$ can be easily computed applying to each state variable the iterations (9)-(11) obtained for the one-factor model.

The extension of the one-factor Beta model to a multiple factors model is not immediate. A rigorous analysis is necessary from both the theoretical and the empirical point of view. This is postponed to future works. Nevertheless, we already obtained results that support the multiple factors extension of the Beta model. Assuming independent factors, we estimated a two-factor Beta model using the extended Kalman filter technique. The univariate threshold structures are estimated as described in the previous section. For both factors we identify three regimes. We do not report the single estimates of the two-factors Beta model since this would be beyond the scope of this section. Nevertheless, in order to support this line of research, we compare the two-factors Beta model with the two-factors CIR model and the two-factor regime switching model of Bansal and Zhou (2002). The corresponding estimated yield curve are plotted in Figure 10 for some given dates, while the pricing errors are reported in Table 5. The two-factors Beta model outperforms both the two alternative model specifications.

[Figure 10 about here.]

Exogenous switching process. The Beta model assumes that conditioned on the current observation of the state variable the regime switching process is determined by Beta distributed shifts defined on multiple thresholds for the state variable(s). Therefore, the model links the current level of the state variable(s) to its (their) future regimes. In this way, the Beta model provides a valuable forecasting tool. Nevertheless, one can argue that the regime classification might be better captured by other indicators than the state variables. In order to incorporate such indicators, the Beta model could be extended to allow a regime switching specification by exogenous (macroeconomic) variables. We briefly describe how this extension can be implemented. For the sake of simplicity we consider a one-factor model. Let X_t be a state variable satisfying the regime switching CIR model of equation (1) and let Z_t be an exogenous macroeconomic variable. The regime switching process s_t satisfies the following dynamics:

$$s_{t+1} = k \Leftrightarrow W_{t+1} \in (w_{k-1}, w_k], \quad \text{for } k = 1, \dots, K, \quad (15)$$

where $w_0 < w_1 < \dots < w_K$ are thresholds for Z_t and the real-valued random variable W_{t+1} is conditionally beta distributed on $[w_0, w_K]$ with parameters $\alpha_{s_t}(Z_t)$ and $\beta_{s_t}(Z_t)$, given the observation of the exogenous macroeconomic variable Z_t and the regime s_t at time t .

If the bond price satisfies the one-factor affine model of equation (7) the one-period no-arbitrage conditions can be derived analogously to the Beta model presented in Section 2. Nevertheless, the probabilities of transition to the next regime are functions of the variable Z_t and thus also the parameters of the affine model.

Applications. Since the proposed Beta model is able to capture both the shape and the level of the yield curve, it can be successfully applied for pricing and risk management purposes. In particular, the

regimes classification through state variables or exogenous macroeconomic variables and their thresholds structures allows to use the Beta model to forecast the yield curve, as already discussed in the previous sections. Future works will be devoted to the pricing of interest-rate-sensitive instruments: swaps, swap-options, Eurodollar futures, T-bond futures, etc.. The risk management implications of our model will be also explored. We expect the Beta model to be a valid instrument for the implementation of risk management strategies.

5 Conclusions

We presented a term structure model allowing for time-varying non-linear conditional first and second moment dynamics of the short-term interest rate process. Similarly to a large number of other studies introduced in the literature, our model belongs to the class of affine term structure models where the dynamics of the short-term interest rate process (or, more generally, of the unobservable factors) are subjected by changes in regimes. In particular, in our model changes in regime occur in response to discrete beta-distributed regime shifts constructed on multiple thresholds.

Estimating the model using an extended linearized Kalman filter on monthly US treasury bond yields at eleven different maturities, we collected empirical evidence that (i) the short-term interest rate dynamics move across more than two regimes, and (ii) our model shows a strong potential in estimating the dependence structure across the different maturities, also in comparison to other standard linear and non-linear affine benchmark models.

We provide some insight into possible future generalization of the proposed term structure model (for example, the incorporation of exogenous macroeconomic variables in the model) as well as possible financial applications where the fit from our model could have a significant impact and allow for better decisions (for example, the valuation of interest-rate-sensitive securities and interest rate management). These and other related issues are currently being explored in ongoing research.

References

- AÏT-SAHALIA, Y. (1996a): “Non-Parametric Pricing of Interest Rate Derivative Securities,” *Econometrica*, 64(3), 527–560.
- (1996b): “Testing Continuous-Time Models of the Spot Interest Rate,” *Review of Financial Studies*, 9(2), 385–426.
- ANG, A., AND G. BEKAERT (2002a): “Regime Switches in Interest Rates,” *Journal of Business and Economic Statistics*, 20(2), 163–182.
- (2002b): “Short Rate Nonlinearities and Regime Switches,” *Journal of Economic Dynamics and Control*, 26(7-8), 1243–1274.
- AUDRINO, F. (2004): “Tree-Structured Multiple Regime in Interest Rates,” Working Paper, University of Lugano.
- AUDRINO, F., AND P. BÜHLMANN (2001): “Tree-Structured GARCH Models,” *Journal of the Royal Statistical Society (Series B)*, 63(4), 727–744.

- BANSAL, R., AND H. ZHOU (2002): “Term Structure of Interest Rates with Regime Shifts,” *Journal of Finance*, 57(5), 1997–2043.
- COX, J., W. INGERSOLL, AND S. ROSS (1985): “A Theory of the Term Structure of Interest Rate,” *Econometrica*, 53(2), 385–407.
- DAI, Q., K. SINGLETON, AND W. YANG (2003): “Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields,” Working Paper, Stanford University.
- DUFFEE, G., AND R. STANTON (2004): “Estimation of Dynamic Term Structure Models,” Working paper, U.C. Berkeley.
- EFRON, B., AND R. TIBSHIRANI (1993): *An Introduction to the Bootstrap*. Chapman & Hall, London.
- FAMA, E. (1984): “Term Premium in Bond Returns,” *Journal of Financial Economics*, 13(4), 529–564.
- FAMA, E., AND R. BLISS (1987): “The Information in Long-Maturity Forward Rates,” *The American Economic Review*, 77(4), 680–692.
- FINK, K. (2004): “A Regime Switching Term Structure Model with Time Varying Transition Probabilities,” Working Paper, James Madison University.
- FRIEDMAN, B. (1984): “Lessons from the 1979-82 Monetary Policy Experiment,” *The American Economic Review*, 74(2), 382–387.
- GALLANT, A., AND G. TAUCHEN (1996): “Which Moments to match?,” *Econometric Theory*, 12(4), 657–681.
- GARCIA, R., AND P. PERRON (1996): “An Analysis of the Real Interest Rate Under Regime Shifts,” *Review of Economics and Statistics*, 78(1), 111–125.
- GRAY, S. (1996): “Modeling the Conditional Distribution of Interest Rates in a Regime-Switching Process,” *Journal of Financial Economics*, 42(1), 27–62.
- LONGSTAFF, F., AND E. SCHWARZ (1992): “Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model,” *Journal of Finance*, 47(4), 1259–1282.
- STANTON, R. (1997): “A Non-Parametric Model for the Term Structure Dynamics and the Market Price of Interest Rate Risk,” *Journal of Finance*, 52(5), 1973–2002.

Notes

¹Let $V \sim \text{Beta}(\alpha, \beta)$ be a Beta distributed random variable on $[a, b]$ for $a < b$, then V has density function $f_{a,b}(x) = \frac{1}{B(\alpha, \beta)(b-a)^{\alpha+\beta-1}} (b-x)^{\beta-1} (x-a)^{\alpha-1}$, where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is the Beta function and $\Gamma(\beta) = \int_0^\infty t^{\beta-1} e^{-t} dt$ is the gamma function. The mean and the variance of V are $\mu = a + (b-a) \frac{\alpha}{\alpha+\beta}$ and $\sigma^2 = \frac{(b-a)^2 \alpha \beta}{(\alpha+\beta)^2 (\alpha+\beta+1)}$, respectively. The mode is at $a + (b-a) \frac{(\alpha-1)}{\alpha+\beta-2}$. The cumulative distribution function corresponds to

$$\mathbb{P}[V \leq x] = \int_a^x f_{a,b}(y) dy \stackrel{z=\frac{y-a}{b-a}}{=} \int_0^{\frac{x-a}{b-a}} f_{0,1}(z) dz = I\left(\frac{x-a}{b-a}; \alpha, \beta\right)$$

where $I(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^x (1-u)^{\beta-1} u^{\alpha-1} du$ is the incomplete beta function.

²The 1 to 6 months discount bond yield has been annualized by multiplying the original data by $\frac{30.4}{365}$.

³The noise term is included in order to avoid singularity of the yield curve variance-covariance matrix when doing the estimation. In fact, the Kalman filter iteration minimizes the pricing error by solving a quadratic optimization problem, thus the inverse of the variance-covariance is needed. Nevertheless, we find that the optimal Kalman filter estimate for the additional ϵ parameter is very small, in the order of 10^{-5} .

Maturity	1m	3m	5m	6m	1y	2y	3y	4y	5y
Mean	5.592	6.012	6.163	6.231	6.437	6.637	6.796	6.915	6.977
Std Dev	2.546	2.671	2.692	2.694	2.629	2.576	2.507	2.480	2.460
Skewness	1.139	1.128	1.111	1.107	1.044	1.011	1.005	0.987	0.965
Kurtosis	4.491	4.399	4.333	4.318	4.123	3.985	3.937	3.876	3.773
Minimum	1.318	1.712	1.758	1.782	1.956	2.368	2.960	3.118	3.346
5% Qtl	2.438	2.753	2.871	2.891	2.999	3.199	3.504	3.628	3.671
10% Qtl	2.845	3.059	3.153	3.239	3.456	3.776	3.929	3.992	4.022
25% Qtl	3.857	4.226	4.397	4.419	4.623	4.830	5.084	5.195	5.239
50% Qtl	5.057	5.402	5.533	5.635	5.911	5.925	6.171	6.548	6.644
75% Qtl	6.743	7.354	7.498	7.540	7.682	7.853	7.967	8.091	8.116
95% Qtl	10.270	11.686	11.815	11.938	11.872	12.219	12.249	12.246	12.249
Maximum	16.146	16.042	16.203	16.526	15.812	15.639	15.557	15.824	15.001

Table 1: Summary statistics of Monthly Yield data (annualized returns in %) for eight maturities. There are 504 observations ranging from January 1960 to December 2001. The data is obtained from CRSP Fama (1984) and Fama and Bliss (1987) files.

Model	Regimes	β_{s_t}	
	s_t	$\beta_{s_t,0}$	$\beta_{s_t,1}$
CIR-Beta Regimes 2 regimes $\beta_{s_t}(X_t) = \beta_{s_t,0} + \beta_{s_t,1} X_t$	1	0	1.643546 (1.954777*)
	2	82.08529 (0.2901734)	-4.766026 (2.285277*)
CIR-Beta Regimes 3 regimes $\beta_{s_t} \equiv const$	1	4.981632 (1.192709)	-
	2	14.15583 (1.874307*)	-
	3	33.4955 (1.988078*)	-

Model	Regimes	mean		volatility
	s_t	$\theta_{s_t}(\%)$	κ_{s_t}	$\sigma_{s_t}(\%)$
Global CIR	1	5.4486333 (4.925250**)	0.0311962 (2.797973**)	0.2746158 (16.700735**)
CIR-Beta Regimes 2 regimes $\beta_{s_t}(X_t) = \beta_{s_t,0} + \beta_{s_t,1} X_t$	1	6.90910399 (2.631956**)	0.008769312 (0.798741)	0.152640449 (8.320622**)
	2	6.22981695 (2.2114569**)	0.05433387 (1.323872)	0.45161741 (6.907943**)
CIR-Beta Regimes 3 regimes $\beta_{s_t} \equiv const$	1	4.686016 (3.239338**)	0.11635685 (2.501473**)	0.1080101 (2.069540*)
	2	5.214959 (2.090928*)	-0.037664065 (1.237196)	0.17421555 (6.683811**)
	3	6.294021 (1.856771*)	0.06241335 (1.469722)	0.4727232 (6.260780**)

Table 2: Beta model estimates, for two different choices of the functions β_{s_t} . The threshold-estimates are $(d_0, d_1, d_2) = (0\%, 8.382\%, 20\%)$ for a 2-regimes model and $(d_0, d_1, d_2, d_3) = (0\%, 3.859\%, 8.382\%, 20\%)$ for a 3-regimes model. The parameter estimates for the global CIR model are also given for comparison. P-values for the null hypothesis that the parameters are equal to zero are given between parentheses. Standard errors are computed using a model-based bootstrap; see, for example, Efron and Tibshirani (1993). One and two asterisks denote significance at the 5% and 1% confidence level.

Model	AIC	drift		volatility		R^2
		MAE (%)	RMSE (%)	MAE (%)	RMSE (%)	
Global CIR	953.60	0.44994	0.72449	0.59432	1.83561	0.12753
CIR regimes switching with constant probabilities	795.75	0.45212	0.72715	0.58501	1.75576	0.20340
CIR regimes switching time-varying probabilities	787.55	0.45091	0.72553	0.58948	1.73671	0.21271
CIR Beta regimes 2 regimes $\beta_{s_t}(X_t) = \beta_{s_t,0} + \beta_{s_t,1} X_t$	815.61	0.44949	0.72218	0.59176	1.68247	0.22826
CIR Beta regimes 3 regimes $\beta_{s_t} \equiv \text{const}$	788.89	0.44062	0.72140	0.58549	1.65925	0.23522

Table 3: Goodness of fit statistics for the short rate. The table shows the AIC statistic for several models for the short rate: the global CIR model, the regimes switching CIR model of Bansal and Zhou (2002) with constant and time-varying probabilities and the Beta regimes models. We also report the in-sample estimation error for conditional first and second moments with respect to the mean absolute error (MAE) and root mean squared error (RMSE) statistics. Additionally, the R^2 statistics is given for the volatility (see Audrino 2004, equation (20) for a mathematical definition of R^2).

Regimes s_t	Beta $\beta_{s_t,0}$	mean		volatility $\sigma_{s_t} (\%)$	price of risk λ_{s_t}
		$\theta_{s_t} (\%)$	κ_{s_t}		
1	4.9454 (0.00012)	-4.2893 (0.00576)	-0.00663 (0.00019)	0.16674 (0.00241)	-0.13045 (0.00480)
2	14.175 (0.00000)	6.2054 (0.00128)	0.00320 (0.00015)	0.13109 (0.00070)	0.41238 (0.00170)
3	33.500 (0.00000)	9.8273 (0.00105)	0.08658 (0.00129)	0.03551 (0.00005)	0.17811 (0.00041)

Table 4: Kalman filter parameter estimates of the 3-regimes constant Beta model fitted on the yield curve for eleven different maturities. The Kalman filter recursions start with the maximum-likelihood estimates from the fit of the univariate short rate model. The regime-dependent market prices of risk are initially set to zero. Standard errors of the parameters are computed using the estimate of the asymptotic variance-covariance matrix based on the outer product of first derivatives of the log-likelihood function (see Duffee and Stanton, 2004) and are shown between parentheses.

	1-factor CIR	1-factor RS	1-factor Beta	2-factor CIR	2-factor RS	2-factor Beta
mean	47	43	41 (36)	30	23	22 (20)
st. dev.	28	27	24 (23)	18	16	16 (15)
min	5	4	4 (4)	3	3	2 (2)
max	174	175	162 (162)	121	114	110 (110)

Table 5: **Model comparison.** Prediction errors (mean absolute error) for several models: one-factor and two-factors regime switching model of Bansal and Zhou (2002) with two-regimes, one-factor and two-factors CIR model and, the one-factor and two-factors 3-regimes constant Beta model. The parameter estimates for the CIR and regime switching models are taken from Bansal and Zhou (2002), thus the prediction errors correspond to those reported in this latter reference. Consequently, for comparison, the sample period for the Beta model is also restricted to that of Bansal and Zhou (2002) and ranges from June 1964 to December 1995, with 379 observations. For the 3-regimes Beta models we also give the prediction error over the whole sample period from January 1960 to December 2001, with 504 observations.

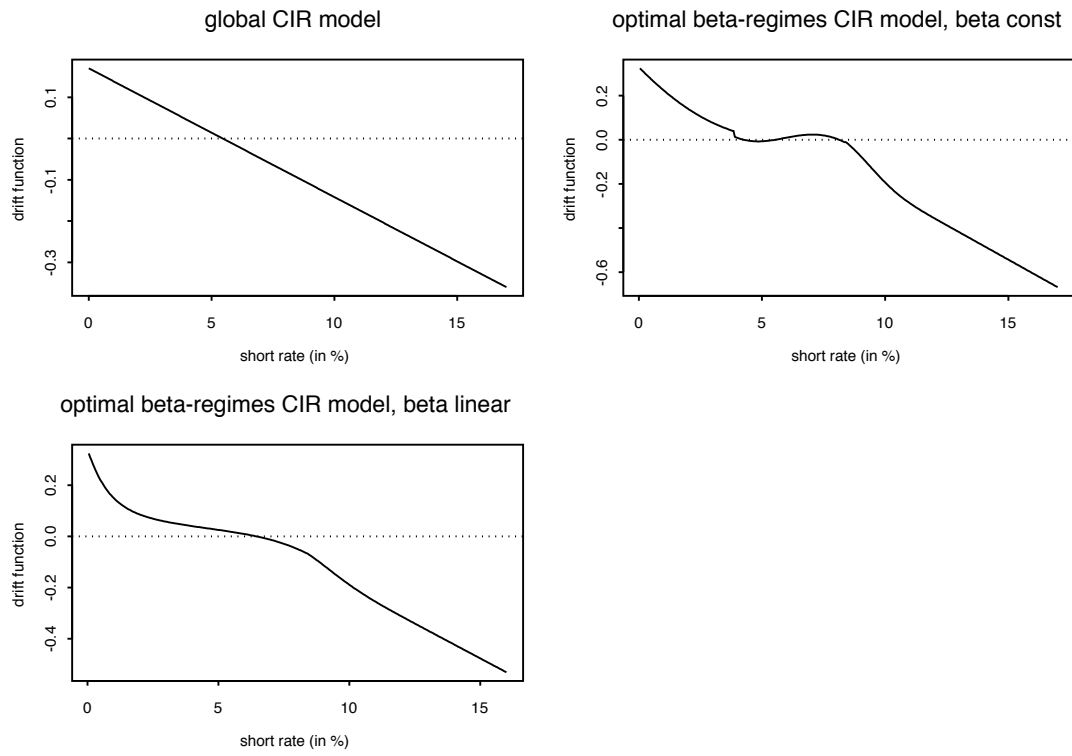


Figure 1: Conditional drift of the short rate for several models: Global CIR model (upper-left panel), Beta model with β_{st} constant (upper-right panel) and Beta model with β_{st} linear in X_t (lower-left panel). Note that the maximum of the short rate in our data set is about 16%.

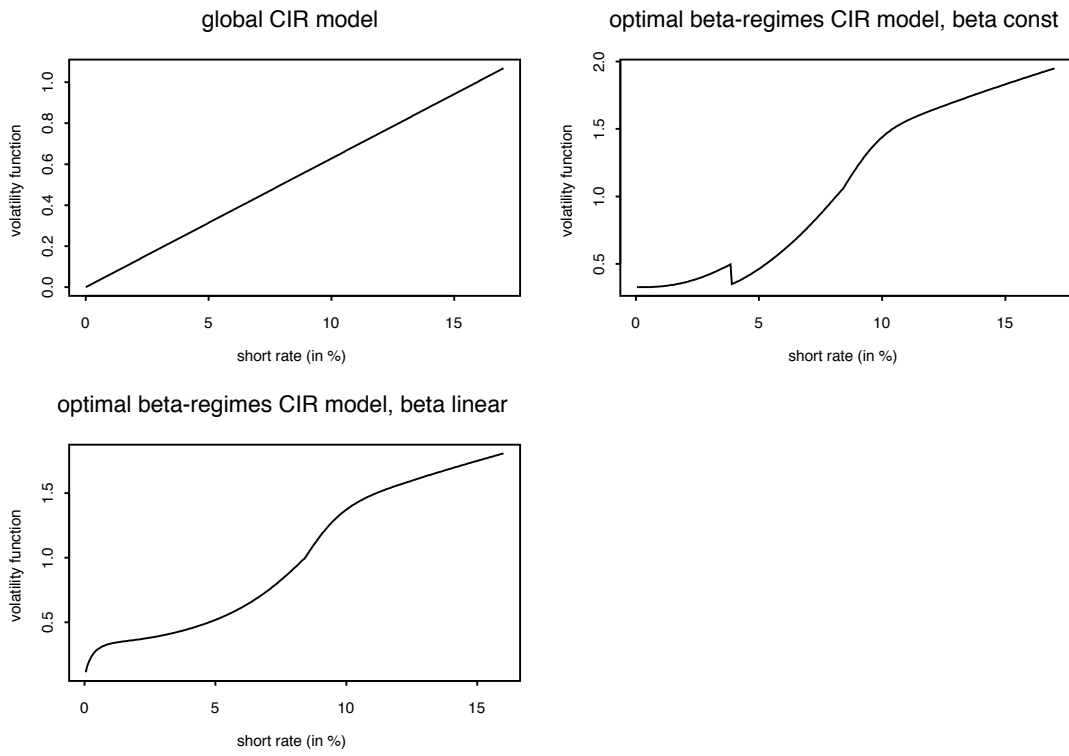


Figure 2: Conditional volatility of the short rate for several models: Global CIR model (upper-left panel), Beta model with β_{st} constant (upper-right panel) and Beta model with β_{st} linear in X_t (lower-left panel). Note that the maximum of the short rate in our data set is about 16%.

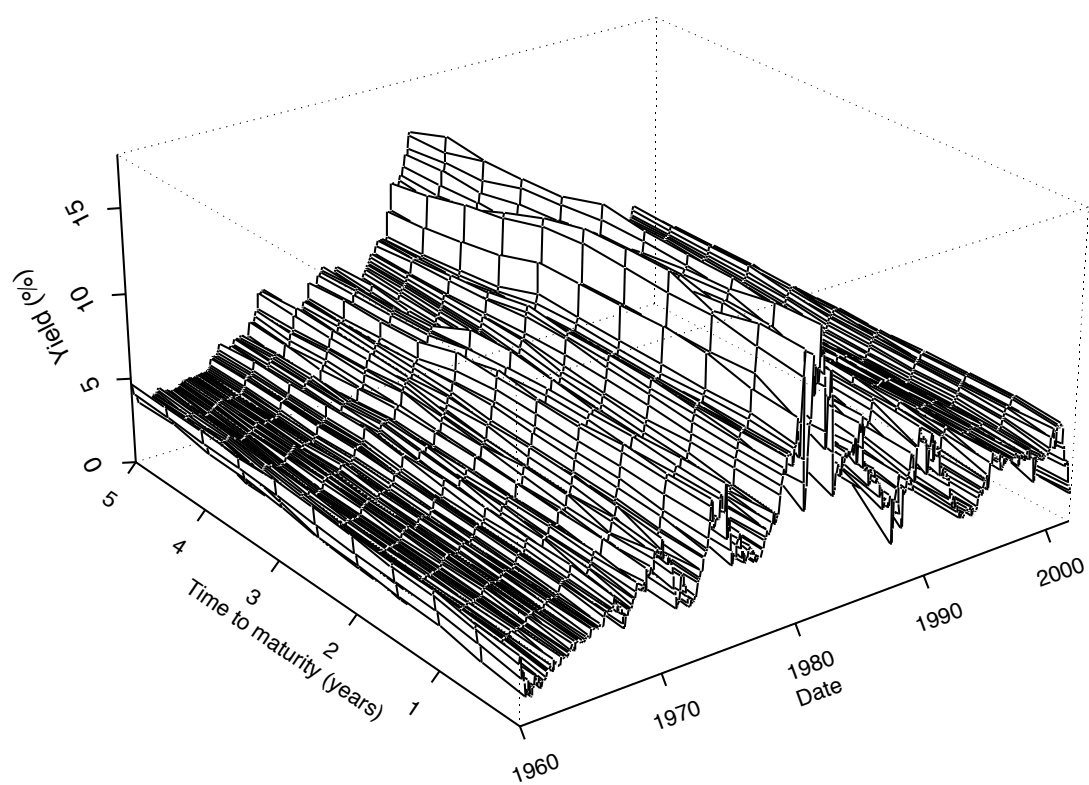


Figure 3: Monthly yield curves over the sample period from January 1960 to December 2001.

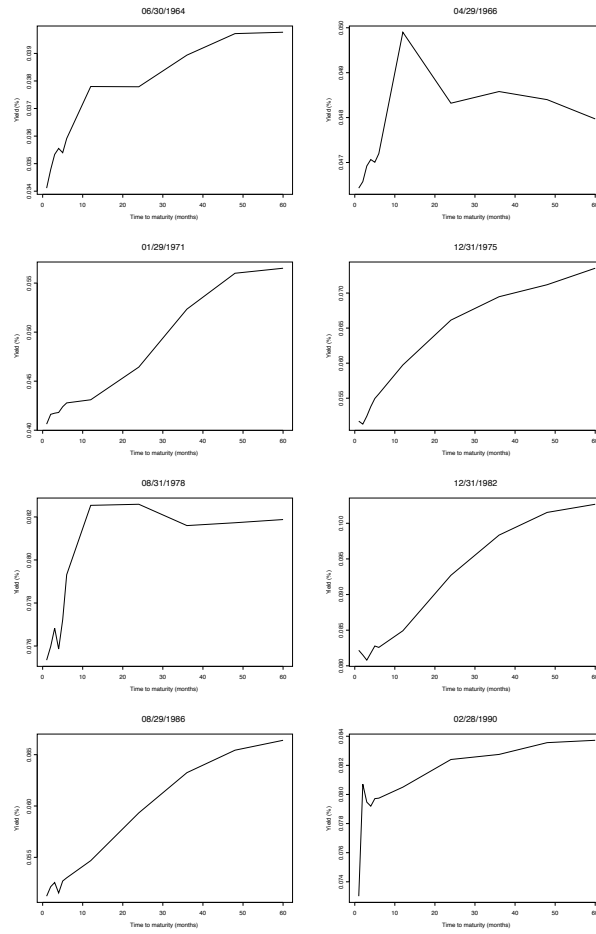


Figure 4: Qualitatively distinct shapes exhibited by realized yield curves for some given dates from the sample period January 1960 to December 2001.

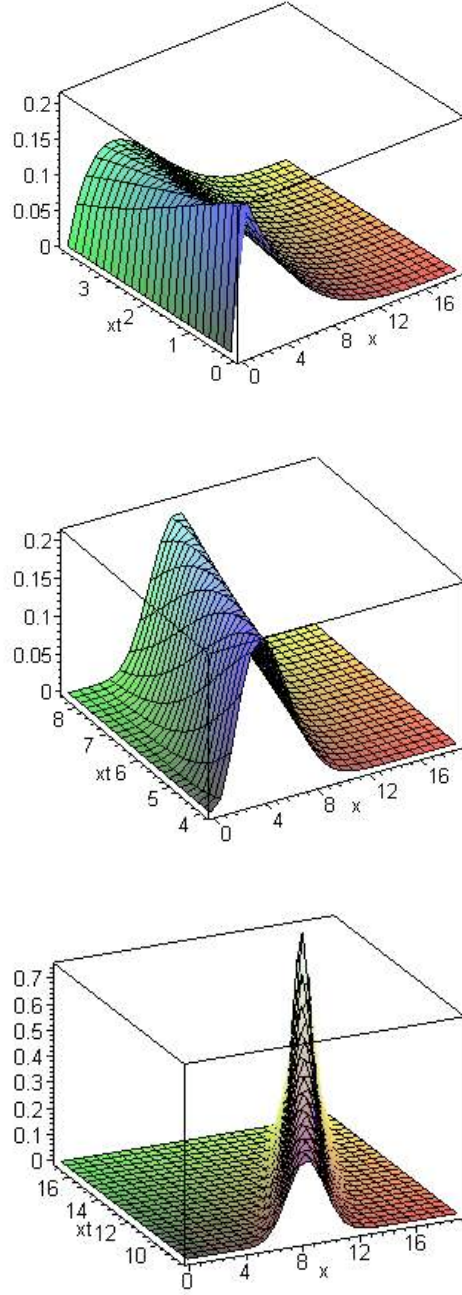


Figure 5: Conditional Beta transition density functions for the three-regimes, constant β_{s_t} model. The x_t -axes gives the current observation of the short-rate, while the x -axes gives the values of the random variable V_{t+1} , ranging from 0% to 17%. The three distributions corresponding to the three regimes are plotted: Regime 1 (upper panel) for $x_t \in [0, 3.859\%)$, regime 2 (middle panel) for $x_t \in [3.859\%, 8.382\%)$ and regime 3 (lower panel) for $x_t \in [8.382\%, 17\%]$.

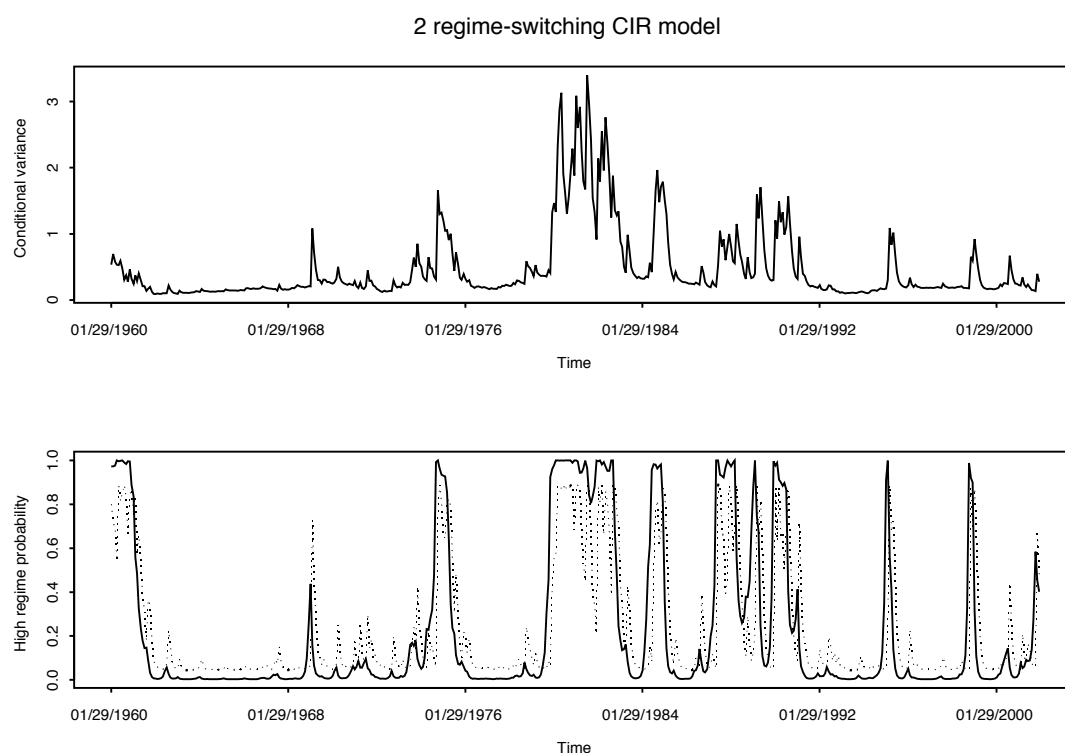


Figure 6: Regime switching model of Bansal and Zhou (2002) with two regimes and time-varying transition probabilities. The upper panel shows the estimated short-rate conditional variance, while the lower panel gives the conditional probabilities for transition to the higher regime.

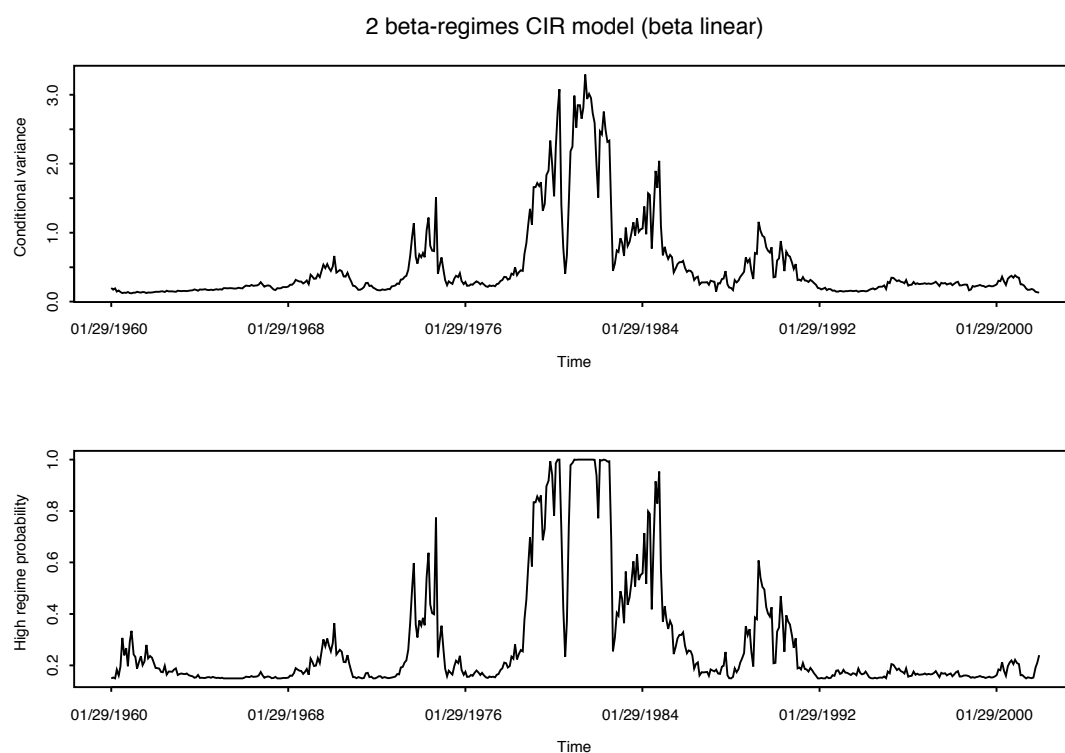


Figure 7: Two-regimes linear Beta model. The upper panel shows the estimated short-rate conditional variance, while the lower panel gives the conditional probabilities for transition to the higher (solid line), middle (dotted line) and low (dot-dashed line) regimes.

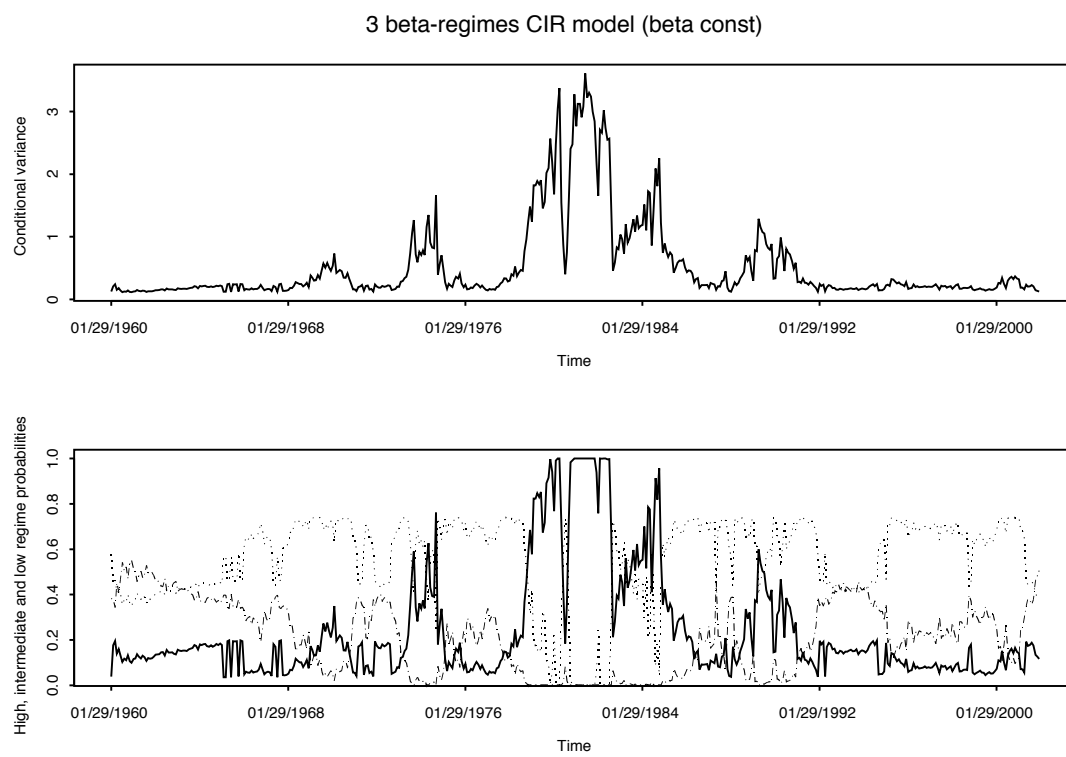


Figure 8: Three-regimes constant Beta model. The upper panel shows the estimated short-rate conditional variance, while the lower panel gives the conditional probabilities for transition to the higher regime.

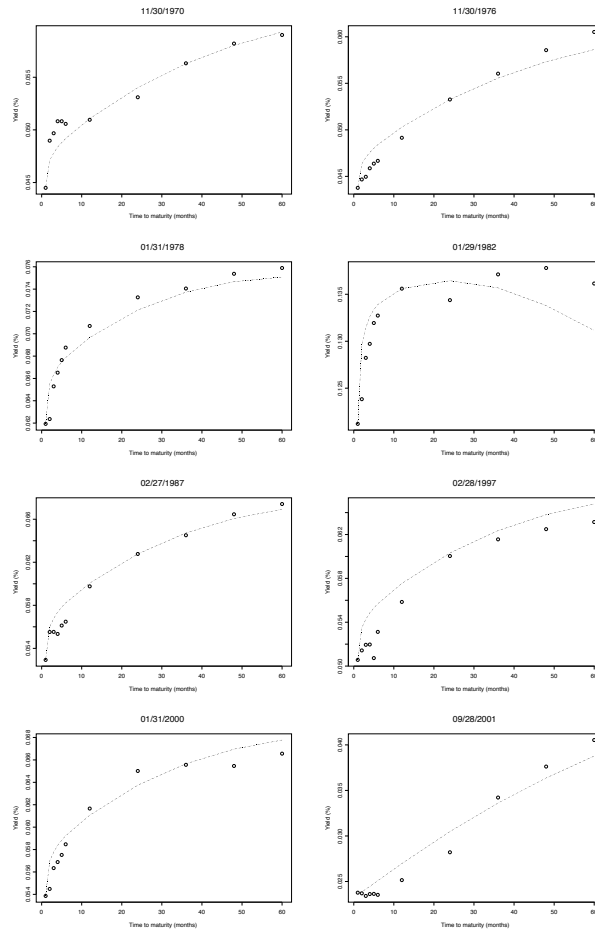


Figure 9: Estimated (lines) and realized (points) yield curve for some given dates in our sample period going from January 1960 to December 2001. The yield curve estimates are from the one-factor 3-regimes constant Beta fit.

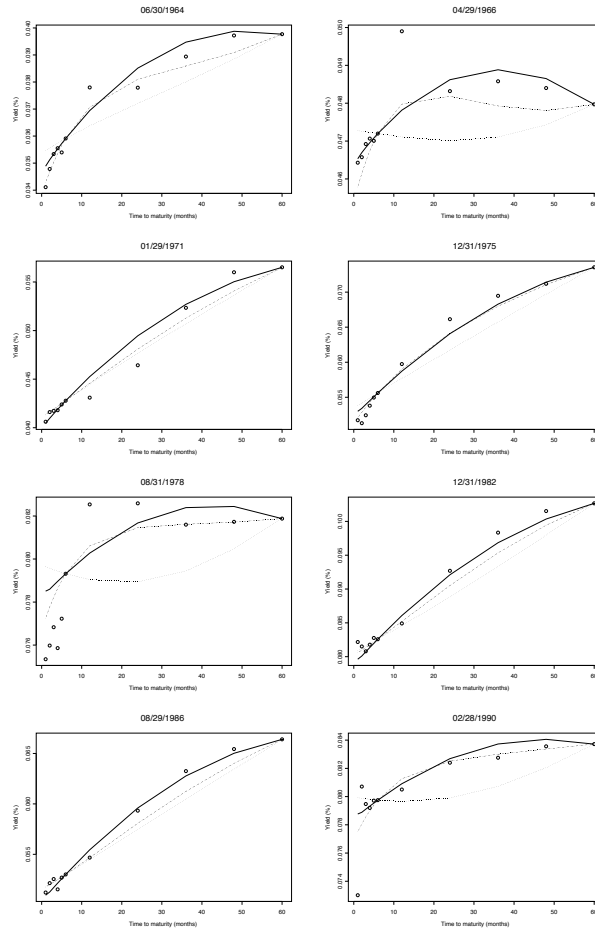


Figure 10: Estimated (lines) and realized (points) yield curve for some given dates in the sub-sample period going from June 1964 to December 1995. The yield curve estimates are from the two-factors 3-regimes constant Beta model (full line), the two-factors regime switching model (dashed line) and the two-factors CIR model (dotted line).